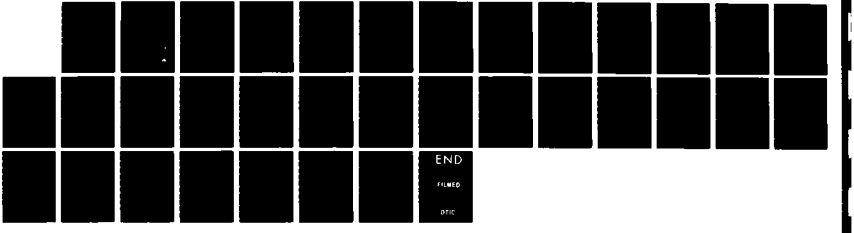
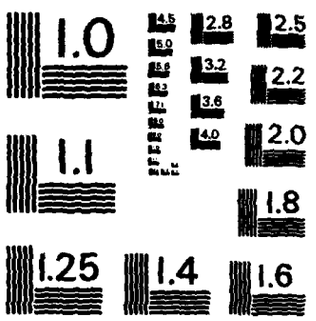
AD-A159 190 SYLLOGISTIC REASONING TASKS A METHODOLOGICAL REVIEW(U) 1/1
SCHOOL OF AEROSPACE MEDICINE BROOKS AFB TX
A NARVAEZ ET AL. JUL 85 USAFSAM-TR-84-41

UNCLASSIFIED F/G 5/10 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

USAFSAM-TR-84-41

SYLLOGISTIC REASONING TASKS

A Methodological Review

Alice Narvaez, M.A.
James C. Miller, Ph.D.

July 1985

Final Report for Period June – August 1983

Approved for public release; distribution is unlimited.

USAF SCHOOL OF AEROSPACE MEDICINE
Aerospace Medical Division (AFSC)
Brooks Air Force Base, TX 78235-5301

NOTICES

This final report was submitted by personnel of the Crew Technology Division, USAF School of Aerospace Medicine, Aerospace Medical Division, AFSC, Brooks Air Force Base, Texas, under job order 2312-V5-27.

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility nor any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise in any manner construed as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this report, and it is releasable to the National Technical Information Service where it will be available to the general public, including foreign nationals.

This report has been reviewed and is approved for publication.

JAMES C. MILLER, Ph.D.
Project Scientist

WILLIAM F. STORM, Ph.D.
Supervisor

ROYCE MOSER, Jr.
Colonel, USAF, MC
Commander

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A159190

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S) USAFSAM-TR-84-41	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION USAF School of Aerospace Medicine 6b. OFFICE SYMBOL (If applicable) USAFSAM/VN	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code) Aerospace Medical Division (AFSC) Brooks Air Force Base, TX 78235-5301	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION USAF School of Aerospace Medicine 8b. OFFICE SYMBOL (If applicable) USAFSAM/VN	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State and ZIP Code) Aerospace Medical Division (AFSC) Brooks Air Force Base, TX 78235-5301	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO. 61102F; PROJECT NO. 2312; TASK NO. V5; WORK UNIT NO. 27
11. TITLE (Include Security Classification) Syllogistic Reasoning Tasks: A Methodological Review	
12. PERSONAL AUTHOR(S) Narvaez, Alice; and Miller, James C.	
13a. TYPE OF REPORT Final 13b. TIME COVERED FROM Jun 83 TO Aug 83	14. DATE OF REPORT (Yr., Mo., Day) 1985 July 15. PAGE COUNT 34
16. SUPPLEMENTARY NOTATION	

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	Cognitive psychology; Reasoning; Performance testing; Syllogisms;
05	10		

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

The syllogism is a system of reasoning both used and investigated in psychology. Its components, defining characteristics, and basic models of form representation are described in this report; and sources of errors in syllogistic problem solving are discussed. Formal rules with which to test the validity of categorical syllogisms are proposed as guidelines for syllogism task construction, and suggestions for experimental design are presented.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS ☐	21 ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL James C. Miller, Ph.D.	22b. TELEPHONE NUMBER (Include Area Code) (512) 536-3464 22c. OFFICE SYMBOL USAFSAM/VN

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

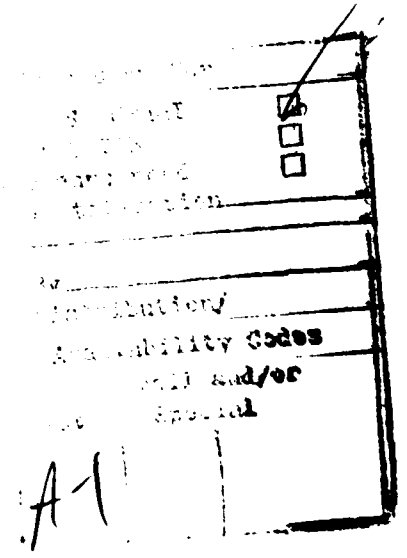
SUMMARY

The syllogism is a system of reasoning that has been both used and investigated in psychology. The syllogism has many components and defining characteristics. Four basic models of form representation (spatial, linguistic, mixed, and algorithmic) used for linear syllogisms are briefly described. Debate concerning the processes that subjects use in solving syllogisms has led to the formulation of process models. ~~Sternberg's (1980)~~ proposed model for linear syllogisms predicts the latency for problem solution as a result of difficulty, while ~~Erickson's (1974)~~ "set analysis" theory for categorical syllogisms provides a way to break down the stages of syllogistic problem solving in order to predict performance. Many factors have been shown to affect or influence subjects' performance in syllogistic reasoning. Two historical hypotheses concerning sources of error are the atmosphere hypothesis--which states that a subject is influenced by the atmosphere, or global impression, of the syllogism--and the conversion hypothesis--which emphasizes illicit processes subjects use in solving syllogisms. Other sources of error include figure effects (syllogisms in certain figures are more difficult than others) and personal bias (one's personal beliefs and attitudes influence performance on syllogisms that are emotion laden or of a controversial nature).

Formal rules with which to test the validity of categorical syllogisms are proposed as guidelines for syllogism task construction, and suggestions for experimental design are presented.

Keywords:

FLD 19



CONTENTS

SYLLOGISTIC REASONING TASKS: A METHODOLOGICAL REVIEW

The purpose of this review is to examine the nature of the syllogism, the investigation of syllogisms, and syllogistic reasoning in psychology. This examination includes the characteristics and fundamental components of syllogisms and their different types. Models concerning the forms of representation that subjects use in solving syllogisms are presented, as are process models describing how syllogisms are solved. Several factors identified as sources of error in syllogistic reasoning are discussed. Finally, several guidelines for constructing syllogisms are proposed, as well as several suggestions for experimental design.

THE SYLLOGISM--DEFINITION, FORM, AND VALIDITY

Definition

Aristotle introduced a system of reasoning called the syllogism--a deductive argument in which a conclusion is inferred from two premises. There are two major types of syllogisms, categorical and linear, based on different systems of logic. The first system of logic, Aristotelean, employs statements concerning class relationships; from these we construct categorical syllogisms. These are conventionally presented as "No X are Y; some Y are Z; therefore some Z are not X." The second system of logic concerns reasoning with transitive relationships, e.g., placing objects in relation to each other on some kind of scale. Such arguments can be presented as "John is taller than Paul; Peter is shorter than Paul; therefore Peter is shorter than John." Presented in this form, transitive inferences are sometimes known as linear syllogisms or three-term series problems. Since most of the research reviewed here involved categorical syllogisms, we will examine them more thoroughly than the linear.

A categorical syllogism is a deductive argument consisting of three categorical propositions that contain exactly three terms, each of which occurs in two of the propositions. A categorical syllogism is said to be in standard form when its premises and conclusion are all standard-form categorical propositions and are arranged in a specified order (Copi, 1972). To understand that order, one must first understand the logician's special names for the terms and premises of categorical syllogisms. (For the remainder of this review, the word "syllogism" will indicate categorical syllogism, in standard form and order, unless otherwise specified.)

The conclusion of a standard-form syllogism is a standard-form categorical proposition that contains two (the major and minor) of the syllogism's three terms. The predicate complement (hereafter called the predicate) of the conclusion is known as the major term of the syllogism; and the subject of the conclusion, as the minor term. Thus in the standard-form syllogism "No X are Y; some Y are Z; therefore some Z are not X," "Z" is the minor term and "X" is the major term. The third term of the syllogism, which does not occur in the conclusion but appears in both premises, is called the middle term. In the example above, "Y" is the middle term. The major and minor terms of a standard-form syllogism occur in different premises; the middle term occurs in both. The premise "No X are Y" contains the major term and is therefore called the major premise; the premise "some Y are Z" contains the minor term and is called the minor premise.

A characteristic of a standard-form syllogism is that the major premise is stated first, the minor premise second, and the conclusion last. It must be noted, however, that the major premise is not defined in terms of its position in the syllogism but as the premise that contains the major term (the predicate of the conclusion). Similarly, the minor premise is not defined in terms of its position in the syllogism but is the premise that contains the minor term (the subject of the conclusion).

Form

The form of a standard-form syllogism is determined by the mood of the categorical propositions it contains and the positioning (figure) of its middle term.

Each proposition can appear in one of the following four moods (each represented by a particular letter):

A)	Universal-Affirmative	(All X are Y)
E)	Universal-Negative	(All X are not Y)
I)	Particular-Affirmative	(Some X are Y)
O)	Particular-Negative	(Some X are not Y)

Therefore the mood of a syllogism is represented by three letters: the first names the mood of the syllogism's major premise; the second, that of the minor premise; and the third, that of the conclusion. In the case of

No X are Y; some Y are Z; therefore some Z are not X

the major premise (No X are Y) is an E proposition, the minor premise (some Y are Z), an I proposition; and the conclusion (some Z are not X), an O proposition. The mood of the syllogism is EIO. The mood of a standard-form syllogism, however, does not completely characterize its form.

Two syllogisms may have the same mood but differ in their forms as a result of their figures--the relative positions of their middle terms. For example, the syllogisms

All Y are Z; some X are Z; therefore some X are Y

and

All Z are Y; some Z are X; therefore some X are Y

are both of mood AII, but they are of different forms. In the first syllogism the middle term is the predicate of both premises; while in the second, the middle term is the subject of both premises.

Four figures are possible for syllogisms: The middle term may be 1) the subject of the major premise and the predicate of the minor premise, or 2) the predicate of both premises, or 3) the subject of both premises, or 4) the predicate of the major premise and the subject of the minor premise. These four possible positions of the middle term constitute Figures 1, 2, 3, and 4 respectively. The following is a schematic array of these figures:

	Figure 1	Figure 2	Figure 3	Figure 4
Major premise	Y-X	X-Y	Y-X	X-Y
Minor premise	Z-Y	Z-Y	Y-Z	Y-Z
Conclusion	Z-X	Z-X	Z-X	Z-X

Thus a complete description of the form of any standard-form syllogism requires that both its mood and figure be named. For example, any syllogism of mood AOO in Figure 2 (named more briefly as AOO-2) will have the form "All X are Y; some Z are not Y; therefore some Z are not X."

Standard syllogisms have many different formats. Since four different moods or propositional statements (A,E,I,O) can occur in both the major and the minor premises, 16 premise combinations are available for a syllogistic task (AA, AE, AI, AO, EA, EE, EI, EO, etc.). Each of these 16 premise combinations can appear in each of the four figures; therefore 64 unique-premise syllogisms are possible. These 64 syllogisms are the ones most commonly used in syllogistic-reasoning research and compose what is referred to as the complete syllogistic task. However, each of these 64 unique-premise syllogisms can conclude with each of the four moods, which amounts to a total of 256 distinct forms that standard-form syllogisms may assume.

Validity

From the point of view of logic, the form of a syllogism is its most important aspect. The validity of a syllogism depends exclusively upon its form (mood + figure) and is completely independent of its specific content or subject matter. A valid syllogism is one in which the conclusion necessarily follows from the premises. An invalid syllogism is one in which the conclusion does not necessarily follow from the premises; therefore the correct deduction will be that no valid propositional conclusion can be drawn about the relationship between its subject and predicate. Of the 64 unique-premise syllogisms, only 19 are valid; 45 are invalid.

With linear syllogisms the terms "determinate" and "indeterminate" are synonymous to "valid" and "invalid" respectively. A linear syllogism is referred to as determinate when it is possible to determine from the premises the relation between each possible pair of terms. An example of a determinate linear syllogism is "Susan is taller than Barbara; Barbara is taller than Paula; who is the tallest?" Not all linear syllogisms are determinate; for example, "Susan is taller than Barbara; Susan is taller than Paula; who is the shortest?" This is an indeterminate linear syllogism because it does not permit inference of the relation between each possible pair of terms.

Now that the components of a syllogism have been presented, a brief description of the syllogistic reasoning task, as it is used currently in psychological research, will be given. The subject (S) is presented with a standard-form syllogism and, on the basis of the information in the two premises, must draw a deductive inference about the relation between the predicate (X) and subject (Z) of the conclusion. The S's response can be made in three ways: a) produce a conclusion believed to be logically valid from relating X to Z, b) select one of four or five given conclusions as the logically valid one; or c) evaluate the validity of a given conclusion. In the latter case, the S is presented with a conclusion and must decide whether it follows deductively from the premises (and thus respond "valid" or "invalid, or "true" or "false"). Performance on the syllogistic reasoning task is typically measured as the number of correct conclusions--produced, chosen, or evaluated. Although most of the research reviewed here used the second method, that of Ss selecting a conclusion from a set of alternatives, there seems to be no inherent reason for the investigators having chosen this method. Subjective and technical factors appear to have determined choice of one method over another.

SYLLOGISMS AND PSYCHOLOGY

As a system of logic, the syllogism has traditionally been used to validate arguments; and the logician, quite appropriately, is more interested in the outcome of the reasoning process than in the process itself. Psychologists, however, are interested in both aspects of reasoning--the process and the outcome--with more emphasis on the former. The categorical syllogism has contributed to our knowledge of inferential processes by providing a well-defined task environment in which such processes may be probed (e.g., Henle, 1962; Wason and Johnson-Laird, 1972). These syllogisms have also found their way into general intelligence tests (e.g., Thurstone, 1938), clinical protocols (e.g., Gottesman and Chapman, 1960; Von Domarus, 1944), evaluations of attitudes and prejudices (e.g., Gorden, 1953; Janis and Frick, 1943; Kaufman and Goldstein, 1967), assessments of belief systems (e.g., McGuire, 1960), and investigations of memory processes (e.g., Erickson, 1972; Frase, 1966). Despite this seemingly widespread use of syllogisms among the different areas of psychology, investigation into the fundamental processes involved in such reasoning has only recently been undertaken (e.g., Erickson, 1974). The advantage of using syllogisms as a task in psychological research, therefore, appears to be twofold: they can serve as a measure of deductive reasoning as well as provide insight into the processes underlying such reasoning.

Linear Syllogistic Representation in Memory

In recent years vigorous debate has arisen regarding S's mental representations of the relations among terms of a linear syllogism. This has led to the formulation of several hypotheses and models concerning the form of representation used by individuals solving linear syllogisms. Four basic models have been proposed. The spatial model (e.g., De Soto et al., 1965; Huttenlocher, 1968) states that information from the two premises of a linear syllogism is integrated and then represented in a spatial array. The linguistic model (e.g., Clark, 1969) denies that the information from the two premises is integrated and claims, rather, that it is represented by deep-structural linguistic propositions. A model that appears to provide a compromise between the previous two is the mixed model (e.g., Sternberg and Weil, 1980), in which information from the two premises is first decoded into a linguistic format and then recoded into a spatial format. Finally, the algorithmic model (e.g., Quinton and Fellows, 1975) asserts that a surface-structural linguistic representation of the premise information is sufficient to solve linear syllogisms and that Ss can use it to bypass the more sophisticated representations proposed by the previous models. For example, when a S is asked "Who is tallest?" a simple set of rules is used to answer the question.

Process Models of Syllogistic Reasoning

Linear

The competing models of form representation have given rise to debate concerning the processes that Ss use in solving linear syllogisms. The investigators who proposed different representations of information have also proposed different processes to use with them. Sternberg (1981) asserts,

however, that the debate regarding representation and process has proceeded with an incomplete data base because indeterminate linear syllogisms, as separate from the determinate, have been neglected in research. To extend the understanding of how people solve indeterminate as well as determinate linear syllogisms, Sternberg (1981) proposed an information-processing model based upon his mixed model of form representation. According to his processing model, indeterminate linear syllogisms are assumed to be easier to solve, on the average, than determinate ones. This assumption is based on the notion that to construct a single three-item array from the two-item arrays, one may need to construct a determinate relation between only two of three possible pairs of relations. On the other hand, a determinate linear syllogism requires construction of a three-item array showing determinate relations between all three possible pairs. However, processing of indeterminate linear syllogisms can be facilitated only if the subject is trained to recognize that such syllogisms are indeterminate. The model assumes that recognition occurs once the individual premises are each linguistically and spatially encoded. These encodings occur regardless of whether the syllogism is determinate or indeterminate.

To validate the proposed mixed model, Sternberg (1981) conducted an experiment in which Ss were asked to solve both determinate and indeterminate linear syllogisms. The latency (reaction time) for solving the syllogisms served as the method of analysis; mathematical modeling of the latency data showed the success of the proposed mixed model in accounting for Ss' performance. However, the fit of the mixed model to the latency data was well below the reliability of those data; the mixed model therefore can be viewed only as an approximation to the strategy that the Ss actually used.

Categorical

The general intent of studies investigating the cognitive processes involved in performing categorical syllogisms has been to clarify the operations that subjects use to interpret premises, combine information, and draw conclusions (e.g., Dickstein, 1975, 1976). Erickson (1974) developed a theoretical model, called a set-analysis theory of syllogistic reasoning, in which syllogistic inference is viewed as a complex process that involves at least three stages: a) interpreting the premises, b) combining the premises after interpretation, and c) choosing a verbal label to describe the interpretation. The hypothesized processes of each stage are illustrated on the next page. In the first stage the premise "All P are M" can mean that P is a subset of M (the circled P enclosed with M) or that P and M are identical (P and M in one circle). In the second stage the relationship of the subset to the superset is indicated. The final stage is the labeling of set relations of S and P. The only sentence common to all combinations of set relations is "Some S are P"; therefore, that is the only valid conclusion.

Experimental support for this theory has been limited to only a few studies, but these data support the idea that syllogistic reasoning is subject to stage analysis and they indicate that a S's conclusion can be predicted on the basis of the Stage I and Stage II thinking.

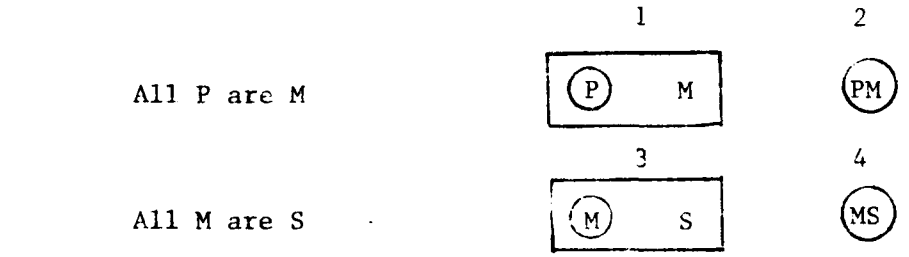
PROCESSES IN DRAWING A CONCLUSION FROM
PREMISES IN A SYLLOGISM (from Erickson, 1974)

Task: Draw a conclusion from premises, or state that no conclusion is logical.

Major premise: All P are M
Minor premise: All M are S
Conclusion: ?

Stage I: Interpretation of premises

Possible Interpretations



Stage II: Combination of interpreted premises

Possible Combinations

Stage III: Labeling of set relation of S to P

Possible Labels

1 and 3
Some S are P, or
Some S are not P

1 and 4
Some S are P, or
Some S are not P

2 and 3
Some S are P, or
Some S are not P

2 and 4
All S are P, or
Some S are P

Logical Conclusion: Some S are P

A variation on Erickson's (1974) set-analysis theory is the transitive-chain model (Guyote and Sternberg, 1978). According to the model, Ss are assumed to encode or interpret the premises of a syllogism in a logically correct and complete manner. Errors in solving the syllogism are thought to occur during the stage in which the premise interpretations are combined. In the example "Some Y are Z; all X are Y," complete encoding of the first premise results in four set relations: Y and Z are equivalent, Y and Z are overlapping, Y is a subset of Z, and Y is a superset of Z. Complete encoding of the second premise results in two set relations: X and Y are equivalent, or X is a subset of Y. There are several possible ways to combine the set relations of the two premises. According to the model, Ss are assumed to combine a maximum of four pairs of presentations, limited by the availability of working memory. The model also assumes that certain set relations are combined before others. For example, identical representations (e.g., X and Y are equivalent) are combined before symmetrical representations (e.g., X and Y are overlapping, and X and Y are disjoint), which are then combined before asymmetrical representations, such as X is a subset of Y and X is a superset of Y. (Symmetrical relation is defined as one in which the positions of X and Y could be reversed without changing the meaning of the representation. In an asymmetrical relation, the positions of X and Y cannot be changed without changing the meaning of the relation.)

Sternberg and Turner (1981) attempted to test the model using a response-evaluation paradigm, as well as to separate experimentally the premise-encoding and premise-combination stages of syllogistic reasoning. They thought that this separation would allow more direct inferences regarding the representations of relations between the X and Z of the premises as encoded and combined by the Ss. Sixty-four Ss were assigned to four groups that differed according to type of task and number of conclusions serving as response alternatives. Subjects in the encoding task group received a premise such as "Some X are not Y" and a conclusion such as "no X are Y." In the combination task, Ss were given a pair of premises such as "Some Y are Z; all Y are X" and a conclusion such as "some X are Z." The Ss' task in both conditions was to indicate whether the conclusion was definitely, possibly, or never true. The subset group consisted of conclusions commonly used in syllogistic reasoning tasks: all X are Z; some X are Z; no X are Z; some X are not Z. The full-set task comprised the subset plus six additional conclusions. The mean proportion of logically correct responses showed a main effect of task, with responses higher for the encoding task than the combination task (across number of conclusions). The model was then tested through model predictions that were derived from parameter estimates in a previous experiment (Guyote and Sternberg, 1978) as well as Sternberg and Turner's current data. The model predicted Ss' responses accurately and further supported the notion that Ss approximate a strategy of complete and correct encoding of the syllogistic premises.

To find out how Ss represent the set relations that characterize each of the premises and the combined premises, Sternberg and Turner (1981) used a modified truth-table analysis. This procedure required the computation of the root-mean-square deviation between the pattern of responses to the 10 conclusions predicted by each of the logically consistent representations and both the observed patterns of responses and the pattern of responses predicted by

the transitive-chain model. From these computations, they could infer for the full-set group (subset plus six additional conclusions) both the representation the Ss actually used for each syllogism and the representation they were alleged to use under the model's assumptions. In fact, all encoded representations were logically complete and correct, as predicted by the model. For example, "All X are Y" was encoded in two ways: as X and Y are coincident and as X is a subset of Y.

Although process models of syllogistic reasoning, such as the ones reviewed here, may seem to focus on differing aspects of the reasoning process, their general intent is to provide a framework with which to understand how syllogisms are solved.

Sources of Error in Syllogistic Reasoning

One common finding that has emerged from the many studies of categorical syllogisms is that Ss frequently derive logically incorrect conclusions. This has led to the formulation of hypotheses and theories that describe and/or explain the factors that appear to influence Ss' performance.

Atmosphere versus Conversion

Historically only two major hypotheses regarding sources of error in syllogistic reasoning have been proposed, and these are offered to account for only a portion of the Ss' errors. These hypotheses are known as the atmosphere effect (e.g., Sells, 1936; Woodworth and Sells, 1935) and the conversion hypothesis (Chapman and Chapman, 1959).

The atmosphere effect, first proposed by Woodworth and Sells (1935), claims that errors in syllogistic reasoning are due to an intrusion of an alogical process--the S judges the validity of conclusions to syllogisms based on a match between a "global" impression formed from the premises with one formed from the conclusion. That is, when the S does not comprehend the relationship between the premises and the conclusion offered, his judgment will be influenced by the atmosphere (global impression) of the premises. According to the hypothesis, the quantity and quality terms of the major and minor premises combine to produce this atmosphere. The quantity of a premise refers to whether the statement is universal or particular. The quality of a premise refers to whether the statement is affirmative or negative. The four combinations of quantity and quality terms yield the four premise types mentioned earlier (A, E, I, O), each with its corresponding atmosphere effect. The atmosphere produced by each of the four propositions are

A = all-yes
E = all-no
I = some-yes
O = some-no

Woodworth and Sells (1935) elaborated on the hypothesis by formulating principles regarding the ways in which these premises combine to influence the deductive process. If both premises are of one kind (A, E, I, or O), atmosphere clearly calls for a similar conclusion. But if one premise is A and the

other I, the atmosphere is partly universal and partly particular, a blend of "all" and "some" which would be weaker than a consistent "all" and thus would amount to a particular conclusion. Similarly, if one premise is E and the other O, the atmosphere is partly affirmative and partly negative, a blend of "are" and "are not," which would be weaker than a consistent "are" and thus would amount to a negative conclusion. Predictions of the conclusions subjects are likely to choose as a result of the atmosphere effect are:

AA = A conclusion
AE, EE = E conclusion
AI, II = I conclusion
AO, EI, EO, IO, OO = O conclusion

Using the example "All X are Y; all Y are Z; therefore all Z are Y," the terms "are" and "all" in the major and minor premises seem to suggest a universal-affirmative atmosphere; and when a S comes to a conclusion that mimics the form of both premises, his or her tendency is to accept it, even though, as in this example, acceptance is incorrect.

Support for the atmosphere hypothesis is provided by Sells (1936). Ninety Ss were presented with 300 syllogisms, each with a single conclusion to be marked as true or false. Significantly more Ss accepted invalid conclusions that were favored by atmosphere than accepted invalid conclusions not favored by atmosphere. Although the atmosphere effect is proposed to account for errors in accepting invalid conclusions, an interesting corollary of the hypothesis appeared relevant to the acceptance of valid conclusions. Since the premises in Sells' syllogisms were structured in correspondence with a set of formal rules (discussed at the end of this paper) to the effect that if one premise is negative, the conclusion must be negative and if one premise is particular, the conclusion must be particular, atmosphere should function to make the valid conclusions less difficult than the invalid ones. In the Sells experiment with 300 syllogisms, the average error rate for the 71 valid items was 16%; for the 229 invalid items, it was 40%.

In 1959, Chapman and Chapman proposed illicit conversion together with probabilistic inference as an alternative to the atmosphere hypothesis. The basic tenet of their hypothesis is that errors are due to a miscoding of propositions and not to faulty inference. Illicit conversion refers to erroneous acceptance of the truth about the converse of a statement. Some conversions such as "No X are Y, no Y are X" and "Some X are Y, some Y are X" are true, but conversions such as "All X are Y, all Y are X" and "Some X are not Y, some Y are not X" are false. Probabilistic inference assumes that if X and Z are both related to Y in a positive way, then they must also be related to each other in a positive way. Similarly, if X and Z are not both related to Y in a positive way, then they must be related to each other in a negative way. Chapman and Chapman maintained that both of these processes result from a reliance by subjects on everyday experience, where acceptance of the converse and probabilistic inference frequently leads to correct conclusions. Chapman and Chapman found that illicit conversion accounted for most of the errors in responses to AA, AE, AI, IA, AO, and OA pairs and that probabilistic inference accounted for most of the errors in responses to II, EE, IE, OO, IO, OI, OE, and EO pairs. Thus, using both principles, Chapman and Chapman predicted that

AA = A conclusion
II, AI, IA = I conclusion
EE, AE, IE = E conclusion
OO, OA, AO, IO, OI = O conclusion
EO, OE = O or E conclusion

When Chapman and Chapman (1959) introduced the principles of illicit conversion and probabilistic inference (the conversion hypothesis) as an alternative to the traditional atmosphere hypothesis, this inspired a line of research dedicated to comparing and contrasting the efficacy of each hypothesis in accounting for the errors that subjects make in solving syllogisms. Simpson and Johnson (1966) claim that there is a considerable overlap in the errors predicted by the two hypotheses; therefore, an experimental manipulation is needed from which different effects on errors can be predicted. Since brief general training reduces errors in logic, Simpson and Johnson reasoned that differential training--one part designed to reduce errors attributable to the atmosphere effect and another to reduce errors attributable to invalid conversion--would serve as such an experimental manipulation. Such a methodology required two blocks, or scales, of syllogisms in which the two error effects operate separately. With syllogisms that confound atmosphere and conversion effects eliminated, five syllogisms (EE, OO, II, IO, EO) remained for a separate test of the atmosphere hypothesis and only one (AO) for invalid conversion.

To test this two-error interpretation, one group of subjects was given a description of the atmosphere effect, a warning against it, and some practice in avoiding it. Another group was given parallel training against the error of invalid conversion. These groups and appropriate control groups were then given both syllogism scales. The atmosphere scale consisted of invalid syllogisms with five alternative conclusions, one of which was an atmosphere error; the five syllogism types were varied in order of terms and propositions so as to make a scale of 12 syllogisms. The conversion scale consisted of 12 AO syllogisms (11 valid syllogisms were included as filler items) varied in order of terms and propositions; with five conclusions, one of which was an invalid conversion.

The results showed that the construction of the two invalid scales was fairly successful: about 70% of the errors on the atmosphere scale were the predicted atmosphere errors, and about 60% of the errors on the conversion scale were the predicted conversion errors. The number correct on the atmosphere scale indicated that the atmosphere-training group made significantly fewer errors than the conversion-training group and each of the control groups. Results for the conversion scale, however, were less clearcut. The number of errors was very large, so an analysis of the number correct was not feasible. Simpson and Johnson (1966) concluded that differential training gave differential results and that neither the atmosphere error nor the conversion error, alone, could account for their findings. The antiatmosphere training had a definitely beneficial effect; the anticonversion-training effect was not as clearcut. Perhaps this difference was due to the quality and execution of the differential training.

In a similar experiment to assess the relative merits of the competing atmosphere and conversion hypotheses, Dickstein (1975) used the approach of differential instructions. The AI (atmosphere instructions) group received special instructions designed to warn them against the error of atmosphere; the CPI (conversion and probabilistic inference) group received special instructions designed to warn against errors of illogical conversion and probabilistic inference; and the third group received standard instructions (SI). Dickstein hypothesized that if the primary cause of error is susceptibility to the atmosphere effect, then performance should be improved by instructions warning against atmosphere errors. Similarly, if the primary causes of error are illogical conversion and probabilistic inference, then the CPI group should do best. Sixty-six Ss performed 64 syllogisms. The data were analyzed by the mean number correct. For valid syllogisms, instructions had no significant effect; but for invalid syllogisms, instructions had a significant effect. Further tests revealed that the CPI group performed significantly better than both the AI and the SI groups, while the SI and AI groups were not significantly different from each other. The finding that instructions against the errors of conversion and probabilistic inference significantly improved performance lends support to the conversion hypotheses proposed by Chapman and Chapman (1959). At the same time, no support emerged for the atmosphere hypothesis; instructions against atmosphere did not improve performance.

Begg and Denny (1969) argued that the explanations of atmosphere effect and of illogical conversion and probabilistic inference differ to such a small degree in the error tendencies they predict that their validity cannot be tested by the error-tendency data. To settle the remaining empirical differences between previous studies, Begg and Denny attempted to gather new data. In the past, atmosphere effects had been defined separately for cases where the two premises of a syllogism were of the same mood (e.g., Woodworth and Sells, 1935) as against cases where the premises differed in mood (e.g., Sells, 1936). Begg and Denny therefore restated the atmosphere hypothesis in terms of two principles that account for and make specific predictions about all cases. The first principle states that whenever the quality of at least one premise is negative, the quality of the most frequently accepted conclusion will be negative; when neither premise is negative, the conclusion will be affirmative. The second principle states that whenever the quantity of at least one premise is particular, the quantity of the most frequently accepted conclusion will be particular; when neither premise is particular, the conclusion will be universal. When these principles are applied to possible premise-pair combinations, the following predictions can be added to those made by Woodworth and Sells (1935) and Sells (1936):

IA = I conclusion
OA, IE, OI, OE = O conclusion

As a result of restating the atmosphere hypothesis, the predictions made by the competing hypotheses (atmosphere vs. conversion) are identical for most pairs, differing only with respect to the IE, EO, and OE pairs. (EI and EA are excluded because they yield valid conclusions in all figures.) The principle of probabilistic inference (conversion hypothesis) predicts an E response to an IE pair, but the principle could equally well predict an O response, congruent with the atmosphere prediction. Begg and Denny (1969) claimed these differences can be reconciled.

Chapman and Chapman (1959) had found predominantly E responses, and Sells (1936) had found O. The question served as a stimulus for Begg and Denny to gather new data. In their experiment (1969), 33 Ss took a 64-item multiple-choice test consisting of four alternative conclusions--one each representing the A, E, I, and O statement moods. The total error response to the invalid syllogisms was calculated for each kind of premise pair. The preferred error tendencies for the pairs were as predicted by Begg and Denny's restated atmosphere hypothesis: When at least one premise was negative, 73% of the errors were negative; when neither premise was negative, 79% of the errors were affirmative. When at least one premise was particular, 90% of the errors were particular; when neither premise was particular, 77% of the errors were universal. These findings illustrate the effectiveness of the atmosphere hypothesis in predicting errors.

The question remains unanswered as to which of the two competing hypotheses, atmosphere or conversion, accounts best for errors subjects make. More recent investigations of the hypotheses have attempted to explain why processes such as illicit conversion occur and to formulate appropriate models to account for error data. One explanation for conversion, suggested by Chapman and Chapman (1959), is that Ss are misled into the particular conversion of universal affirmative propositions by their similarity to definitional statements. For example, a proposition such as "All X are Y" resembles a definitional statement such as "All right angles are 90-degree angles." Because of this similarity, they argued, Ss treat syllogistic propositions as if they were definitional statements, and this leads to an identity interpretation.

More recently Dickstein (1981) proposed that the validity of this explanation be tested by contrasting the traditional format of syllogistic propositions with an alternative format containing modified propositions that still allow identity interpretations but are dissimilar from definitional statements. Such a format might be constructed by modifying syllogisms to emphasize the meaning of the premises by statements about "class membership." For example, the premise "All X are Y" might be rewritten as "All members of class X are members of class Y," thus making it clear that there are two distinct classes and that the issue concerns the relationship between membership in one class and membership in the other. Dickstein reasoned that if similarity to definitional statements is responsible for the illicit conversion of universal affirmative propositions, then less conversion should occur with the modified premises.

Dickstein (1981) also noted that conversion may not be the only factor responsible for errors. Chapman and Chapman (1959) reported that conversion could account for errors on only 13 of the 45 invalid syllogisms included in the traditional syllogism task. On the remaining 32 invalid syllogisms, Ss often drew propositional conclusions even though such conclusions were unjustified. Those 32 syllogisms were labeled invalid nonconversion and have been demonstrated to be consistently less difficult than the invalid conversion syllogisms (e.g., Dickstein, 1975, 1976). Dickstein proposed that one explanation for errors on invalid nonconversion syllogisms is that Ss fail to make the important distinction between "necessary" and "possible" conclusions. Subjects may not realize that a conclusion which is merely possible--but not

demanded--by the premises is not acceptable under the rules of deductive reasoning. This unfamiliarity with the rules may lead Ss to use a less rigorous criterion of evidence when accepting a conclusion as valid. Dickstein hypothesized that Ss who are instructed in this feature of deductive reasoning should perform better on invalid nonconversion syllogisms than Ss who are not given instruction.

In an experiment conducted by Dickstein (1981), Ss were assigned to groups as a result of premise format: a) traditional premises written in classical syllogistic form, and b) instructions clarifying that a propositional conclusion could be drawn only if the conclusion was compelled by information in the premises. Subjects received 64 syllogisms and were required to select one of five responses. Of the 64 syllogisms, 19 were valid; the 45 invalid consisted of 13 conversion and 32 nonconversion syllogisms. Group means for the percentage correct were calculated for valid, invalid conversion, and invalid nonconversion syllogisms. There was a main effect of syllogism type: performance was best on valid, then invalid nonconversion, and then invalid conversion syllogisms. Also, instructions and syllogism type showed interaction. On valid syllogisms, Ss performed better without additional instructions than with; whereas on invalid--conversion and nonconversion--syllogisms, Ss performed better with additional instructions than without. Premise type, however, had no main effect.

The first conclusion drawn from this experiment was that clarifying the role of "necessity" in deductive reasoning significantly improved performance, especially on invalid nonconversion syllogisms. The results supported the argument that poor performance on invalid syllogisms reflects the failure of Ss to confine their inferences to propositions that are compelled by the premises. The second conclusion was that altering the form of the premise does not improve performance; thus this finding casts doubt on the explanation for conversion by Chapman and Chapman (1959). And finally, the results supported the differentiation of two subsets of invalid syllogisms, conversion and nonconversion, in that the former are significantly more difficult than the latter.

Revlis (1975) formulated a process model of reasoning that attempts to predict both reasoning errors and successes. An implicit assumption of the model is that conversion automatically occurs as quantified relations are comprehended. It is not simply one way in which a S might resolve the apparent ambiguity of quantified statements, rather it is the basic reading given to such statements. An exception to this rule occurs only in instances where the pragmatic rules of English prevent the converted interpretation of a sentence because it would be semantically deviant. This conversion model claims that Ss make reasoning errors primarily in cases where the encoding of premises entails converting the problem to one in which a different conclusion is required. It predicts that Ss will be correct when either a) the problem is converted but the conclusion is the same in the converted and original forms of the problem, or b) the S's knowledge of the world blocks conversion. Therefore, Ss make errors in reasoning when their encoding of the premises transforms the syllogism into another problem with a logical conclusion different from the original.

To test the validity of the Revlis conversion model, Revlin et al. (1978) conducted an experiment in which Ss solved 32 syllogisms, half of which were valid and half invalid. The premises were constructed as one universal assertion and one arbitrary relation, and conclusions expressed a relation with no real-world truth or validity. Four types of syllogisms were represented: VALID SAMES and INVALID SAMES--conversion results in a syllogism that has the same conclusion as the presented problem; VALID DIFFERENTS and INVALID DIFFERENTS--conversion produces a syllogism with a different conclusion than the one presented. Subjects were assigned to groups designated by the kind of problems they received: CN--problems in which conversion would produce an acceptable interpretation because the universal sentence and its converse were both true. (Example: All veterinarians are animal doctors; some animal doctors are members of the Finville Club.) IN--problems whose universal sentence expressed an arbitrary inclusion relation, based on which there is no a priori reason to assume that Ss would block conversion. (Example: All products on the list are types of tools; some types of tools in the store are imported objects.) CB--problems on which conversion would be blocked. Example: All crowbars are tools; some tools in the store are imported objects.)

Revlin et al. (1978) hypothesized that when Ss convert premises they reach decisions quite different from when they do not. That is, when conversion is not blocked, Ss should make errors in solving DIFFERENTS problems but perform accurately in solving SAMES problems. When conversion is blocked, however, Ss will be more accurate on DIFFERENTS problems--valid and invalid --and should have no change in reasoning accuracy for SAMES problems--valid and invalid. The accuracy score of each S was the sum of the percentage correct for each problem type. A comparison of the CB and CN groups showed that Ss were more accurate on blocked than unblocked problems (45.9 versus 33.8). As predicted, this was due to a significant interaction between the type of relation (blocked-unblocked) and the type of problem (SAMES-DIFFERENTS). That is, the two groups differed only on problems that the conversion model predicts are affected by conversion--the DIFFERENTS syllogisms. A comparison of the CB and IN groups showed, again, that Ss were more accurate in solving blocked than unblocked syllogisms. This blocked effect was also due to an interaction between relation and problem type in which the major difference between the CB and IN groups was on problems where the conversion model predicts that the encoding of the premises is critical--valid DIFFERENTS syllogisms.

In addition to the specific predictions tested and confirmed by Revlin et al. (1978), very general statements about the relative reasoning level among all groups were also confirmed. For valid SAMES, the model correctly predicted that Ss would show the same level of accuracy across groups. The model also correctly predicted that reasoning accuracy would not differ across the convertible groups (CN and IN) on valid DIFFERENTS (0% correct was predicted). Finally, the model correctly predicted that when a blocked group was included in the analysis (CB), there would be a difference across groups. These results, therefore, lend support to the conversion model by illustrating that when conversion is allowed to enter into the encoding of quantified relations, Ss exhibit the error pattern predicted by the model.

Figure Effects

The figure of a syllogism has also been considered to influence Ss' performance on reasoning tasks. By considering the structural characteristics of standard-form syllogisms, Frase (1968) discovered that syllogistic features bear an exact correspondence to three-stage mediational paradigms used in pair-associate learning. By viewing the syllogism as an analogous device for mediated association, we see that Figure 1 syllogisms (All Y are Z; all X are Y; therefore all X are Z) are analogous to the forward chain as analyzed in mediational terms. For example, in a paired-associates task, a S may be required to associate Y-Z, then X-Y, and finally, in the test condition, to associate X-Z. The X-Z association is facilitated because both X and Z have been associated with the common mediating event, Y. The sequence of associated events set up during learning in the present example can be represented as a forward chain: X-Y-Z. In syllogistic reasoning, the first premise of this example represents learning the Y-Z association; and the second premise, the X-Y relation. After reading the conclusion (the X-Z association) the S must judge whether that association follows from the associations established in the two premises.

Frase (1968) reasoned that mediation paradigms that differentially facilitate paired-associate learning should similarly facilitate logical judgments. The four figures of syllogisms are viewed by Frase as corresponding to the mediation paradigms of forward chain, stimulus equivalence, response equivalence, and reverse chain respectively. Figure 4 syllogisms (All Z are Y; all Y are X; therefore all X are Z) correspond to the reverse-chain paradigm, Z-Y-X. Facilitation should be least strong with this paradigm (e.g., Jenkins, 1963).

To test whether the mediational model can be used to predict the number of errors made in judging the validity of syllogisms, Frase (1968) conducted an experiment in which two groups of Ss (trained and untrained in syllogistic reasoning) were administered a test of 22 syllogisms. Eleven syllogisms had equivalent validity in all four syllogistic figures and violated similar logical rules across the four figures, and 11 valid syllogisms were used as filler items. The Ss' task was to record whether the syllogism was valid or invalid and to record the confidence level of their judgment on a 5-point scale. Figure 4 syllogisms were predicted to produce the greatest number of errors, and Figure 1 syllogisms to produce the least. The means for errors and confidence were calculated for each figure as well as for each group. With regard to error scores, the trained group performed significantly better in judging the validity of syllogisms. The four figures also showed a significant difference. For the confidence of judgments, there was an interaction of figure by group. Subjects in the trained group tended to be more sure of their responses overall and to be least sure of their judgments on Figure 4 syllogisms. The trained-group confidence ratings thus corresponded more closely to the distribution of error scores than did the ratings of the untrained group. Even though the Ss in the trained group were aware that Figure 4 was more difficult than other figures, they tended to make more errors on syllogisms in that figure than on any other.

The major conclusions reached on the basis of this experiment were a) the number of errors made on syllogisms similar to the reverse-chain paradigm was higher than with other paradigms, and b) Figure 1 syllogisms, or the forward-chain paradigm, facilitated reasoning most. Also, Ss' reports on how sure they were of their judgments on the different figures were differentially affected by whether or not they had received training, but error scores were not. In terms of confidence being matched to the relative distribution of errors, the trained group was more successful.

Roberge (1971) claimed that the methodology used by Frase (1968) may have permitted only a conservative test of the mediational model's predictions and may have masked the true differences. For example, Ss may have used nonlogical methods, such as the atmosphere effect, to evaluate conclusions as valid or invalid. Even reasoning illogically, Ss could have evaluated correctly. Roberge attempted to reexamine the effect of mediated associations in deductive reasoning by introducing some potentially important methodological changes: a) Require Ss to select the correct conclusion from among five alternatives, b) assure that the correct conclusion for most items was not the atmosphere effect, and c) include test items for syllogistic moods and figures that were ignored in previous studies.

Fifty-two Ss received a 59-item multiple-choice test consisting of 45 invalid and 14 valid items. Only nine premise pairs were used; this resulted in conclusions having equivalent validity in all four syllogistic figures and also violated similar logical rules across the four figures. Subjects were asked to indicate (5-point scale) how certain they were of their judgment for each syllogism. Mean error scores for the four figures differed significantly. The findings of Roberge (1971) that Ss made more errors on Figure 4 syllogisms and that Figure 1 syllogisms were easier than the Figure 4, confirmed the results reported by Frase (1968). Syllogism figure also had an effect on the confidence level: Pair-wise comparisons indicated a significant difference between mean confidence ratings for Figures 2 and 4. The mean confidence ratings for these two figures corresponded to their mean error scores; Ss were more certain of judgments on the easiest syllogism (Figure 2) and less certain on the most difficult syllogism (Figure 4). These latter findings are not consistent with those of Frase.

Dickstein (1978) argued that, although the analogy between paired-associate learning and syllogistic reasoning is interesting, the mediation analogy cannot be regarded as an adequate explanation for figure differences and their effect on performance. In accordance with the view of syllogistic reasoning as an information-processing activity, Dickstein suggested an alternate explanation for figure effects. He proposed that figure effects occur for syllogisms in which conclusions that are logically justified from Z to X are different from conclusions that are justified from X to Z. Even though the syllogistic task requires Ss to draw conclusions that are appropriate from X to Z, some Ss will draw conclusions that are appropriate from Z to X and then erroneously apply them to the relation from X to Z. Dickstein further proposed that how a S processes the information in the premises--in a forward (X to Z) or backward (Z to X) direction--will be influenced by the direction of the information presented in the premises.

In Figure 1 both premises proceed in a forward direction (X to Z), and Ss should have little tendency to process information from Z to X. On the other hand, in Figure 4 both premises proceed from Z to X; here the tendency to process information in a backward direction should be maximal. Figures 2 and 3 are indeterminate: in both, one premise proceeds in a forward direction and the other proceeds in a backward direction. Therefore, Dickstein (1978) hypothesized that for premise combinations in which forward and backward processing lead to different conclusions, performance will be best on Figure 1 syllogisms, intermediate on Figures 2 and 3, and poorest on Figure 4. Furthermore, this decrement in the number of correct conclusions across figures would be paralleled by an increase in the number of errors resulting from backward processing.

In an experiment to test these hypotheses, Dickstein (1978) randomly divided Ss into four groups. Each group completed all the unique-premise syllogisms for a single figure, with two presentations of each syllogism--for a total of 32 syllogisms. An analysis of variance was conducted for both the number correct and the predicted error. For the correct response, the figure had a significant effect: Performance on Figure 1 was best; on Figures 2 and 3 combined, intermediate; and on Figure 4 poorest. The least error was in Figure 1; an intermediate amount, in Figures 2 and 3 combined; and the greatest amount, in Figure 4. These results indicate that figure is a significant determinant of performance on syllogisms. The data also support the explanation of figure effects in terms of backward processing; the decrement in the number of correct responses across figures was paralleled by an increase in the occurrence of the error that was predicted on the basis of backward processing. The more strongly the premise-information direction promoted backward processing, the greater the number of errors--corresponding to Dickstein's hypothesis that backward processing is done at the expense of a correct conclusion. The analyses of premise combinations for which both forward and backward processing produced the same conclusions gave additional strong support for the backward-processing explanation. In all of these latter instances, there were no significant differences between figures. Thus, figure differences occurred only where the conclusions derived from forward and backward processing were discrepant.

The effects of syllogism figure have been demonstrated to account for particular errors Ss make in syllogistic reasoning. However, neither the mediational model (Frase, 1968; Roberge, 1971) nor the one proposed by Dickstein (1978) are comprehensive and complete in predicting all errors due to figure effects. Perhaps incorporating these effects and, specifically, the principle of backward processing into a model such as Revlis's (1975) conversion model would allow maximal prediction for Ss' performance. The effects of backward processing may be viewed as analogous to the process of illicit conversion that is proposed in Revlis's model. Illicit conversion occurs when Ss erroneously convert premises such as "Some Z are not X" to "Some X are not Z" when encoding the syllogism. Similarly, backward processing occurs when Ss draw conclusions that are appropriate when processing a syllogism from Z to X but erroneously apply them to a syllogism whose relation is from X to Z. In this way, backward processing can be viewed as a new kind of illicit conversion. Backward processing also seems to account well for error tendencies not accounted for by the conversion model.

Personal Bias

A widely held belief is that people are likely to be satisfied with unsound arguments if they accept the conclusion to which such arguments lead, and conversely that they are likely to be unduly critical of sound arguments if they reject the conclusion. Observations such as these have led researchers to question whether logical reasoning is affected by a S's personal beliefs and attitudes. That is, when solving syllogisms of personal relevance, do Ss base their judgments on the believability of the conclusions rather than on the logical form of the argument?

Janis and Frick (1943) proposed two hypotheses regarding personal bias and errors on performance: a) If there is agreement with the conclusion, more errors will be made by accepting invalid arguments than by rejecting valid arguments, and b) If there is disagreement with the conclusion, more errors will be made by rejecting valid arguments than by accepting invalid arguments. Janis and Frick conducted an experiment, based on those hypotheses, to test two propositions: a) Subjects will make more errors on "agree-invalid" items than on "agree-valid" items, and b) Subjects will make more errors on "disagree-valid" items than on "disagree-invalid" items. ("Agree" and "disagree" refer to a S's attitude toward the conclusion as an isolated statement; "valid" and "invalid" refer to formal correctness of the conclusion following from the premises.) Subjects received 16 syllogisms; half were valid and half were invalid. They contained conclusions that were anticipated to produce either general agreement or disagreement, with a minimum of neutral responses, thus allowing fairly equal distribution of items in the following four categories: a) A-V, b) A-I, c) D-V, d) D-I. An example of a syllogism that would fall under the D-V (disagree-valid) group is "All poisonous things are bitter; arsenic is not bitter; therefore arsenic is not poisonous." After performing the syllogism task, Ss were given an attitude test in which the conclusions of all the syllogisms were presented in random order, and then were asked to record whether they disagreed or agreed with each statement.

Response frequencies were tabulated in the four categories. Out of a total of 304 responses on the syllogism task, 70 were wrong. A chi-square test was applied and a significant difference found in that errors had a greater tendency to fall in certain categories than in others. To test whether the results confirmed one or both of the propositions, a chi-square test was applied separately to items for which there was agreement with the conclusion and to items for which there was disagreement. A significant difference was found for both tendencies, thus supporting the proposition that attitudes toward the conclusion are related to errors in judging the logical validity of syllogisms.

To discern whether it is the S's inability to reason correctly or a willful attempt to conform a conclusion to one's personal beliefs that is responsible for errors, Morgan and Morton (1944) conducted several experiments. The first was designed to check the atmosphere hypothesis (Woodworth and Sells, 1935) and to investigate whether the accepted conclusion would be influenced by the type of term used in the propositions. Subjects received a multiple-choice test of 64 syllogisms. Thirty-two syllogisms contained the symbolic terms "X," "Y," and "Z"; the other 32 consisted of the verbal terms "troubles," "unpleasant," and "insults." These words were selected because

they might have some emotional value but probably would not touch off any personal convictions held by the subjects. A chi-square test was used to test the significance of the differences between the conclusions accepted when the terms were abstract symbols and when they were words. Of the 32 syllogisms, 18 had significantly different conclusions from those expected on the basis of logical processes. So that any atmosphere effect could be determined, the data were reduced to percentages and calculated to show the relative part played by atmosphere, correct reasoning, and chance elements. Logic played a small part in the selection of conclusions, as compared to the part played by the atmosphere effect (32.5 versus 46.1 respectively). Morgan and Morton concluded that most Ss do not answer syllogisms by reasoning clearly but are influenced by vague feelings they derive from the form in which the premises are stated. They therefore reasoned that using terms which touch off emotional prejudices in the Ss will accentuate distortion.

In a second experiment addressing this problem, issues concerning World War II were incorporated into 15 syllogisms and given to Ss, together with 15 syllogisms of parallel form but with symbolic premise terms. An example of the former is given below:

> Usually extremely brutal repressive measures are not justified in subjugating a country, but some extremely brutal repressive measures were taken by the Nazis when they killed all of the male inhabitants of the town of Lidice.
>
> 1. The Nazis were justified in killing the male inhabitants of Lidice in subjugating the country.
>
> 2. The Nazis may have been justified in killing the male inhabitants of Lidice in subjugating the country.
>
> 3. The Nazis were not justified in killing the male inhabitants of Lidice in subjugating the country.
>
> 4. The Nazis may not have been justified in killing the male inhabitants of Lidice in subjugating the country.
>
> 5. None of the given conclusions seems to follow logically.

The frequencies for each of the five response alternatives were calculated for the symbolic and verbal syllogisms. The results of a chi-square test showed significant shifts in conclusions selected because of personal convictions. That is, for most of the syllogisms, the distributions of answers that had been given to the symbolic terms changed after the popular issue was introduced, and this shift was in the direction of popular opinion. So that the relative part played by the different factors involved could be determined, the number of individuals who answered each syllogism in accordance with atmosphere effect, logical inference, popular opinion, and other conclusions which reflect none of these factors were calculated.

Morgan and Morton (1944) generally concluded that when syllogisms contain nothing to arouse a response based on personal convictions, the selection of a conclusion is determined about 50% by atmosphere effects and 25% each by logic and chance factors. When an issue is injected that relates to the personal opinions, wishes, fears, or convictions of a S, response selection shifts from the atmosphere effect to the meaning involved in the terms of the syllogism. In response to such syllogisms, atmosphere effect contributes about 25% of the deciding influence; logic and chance factors, each about 20%; and the personal convictions of the respondent, about 35%.

A similar experiment was conducted by Gorden (1953) to determine the extent to which a person's attitude on a subject may interfere with his ability to think logically on that subject. Subjects performed a multiple-choice test of 24 syllogisms: 12 were on the USSR; the other 12 were items of nonsocial and noncontroversial nature. An example of a syllogism on Russia follows:

> Government control and ownership of industry often lead to a more efficient economy by eliminating duplication of functions. Since Russia has more government ownership and control of industry than the United States,
>
> 1. Russia has a more efficient economy than the United States.
>
> 2. Russia may have a more efficient economy than the United States.
>
> 3. Russia does not have a more efficient economy than the United States.
>
> 4. Russia may not have a more efficient economy than the United States.
>
> 5. None of the above conclusions are necessarily logical.

For each of the USSR syllogisms, the conclusions were designed so that two were pro-USSR and two were anti-USSR. In half of the USSR items, the major and minor premises apparently supported the pro-USSR conclusion, and in half they apparently supported the anti-USSR conclusion. However, in each of the 24 items there was some qualification that technically rendered the syllogism invalid so that conclusion number 5 was appropriate. Gorden (1953) hypothesized that a S who is an intelligent reader, makes a cautious commonsense attempt to be logical, and is not influenced by any strong emotional reaction to the content but at the same time does not know the technical points of logic, will choose number 2 or 4 conclusions--50% of each. But a S who has a solid knowledge of logic and is not influenced by a strong emotional reaction, will choose the number 5 conclusion.

Tests of these hypotheses were constructed so that a S's choice of a conclusion other than number 5 could be determined with reasonable certainty as being due to a lack of logic, or subjective emotional responses, or poor reading ability. The syllogisms were conceived in groups of three labeled triplets. In the first item of a triplet, the atmosphere was slanted in a

pro-USSR direction; that is, the major and minor premises would logically support pro-USSR conclusions if it were not for a technical loophole that rendered none of the first four conclusions valid. The second item was a nonsocial issue with the same degree of sentence complexity as the first item. The third item had an anti-USSR atmosphere, with sentence structure parallel to that of the previous two items. In this way the middle item acted as a control on the S's ability to think logically and on reading ability. Between each triplet was interspersed another non-social-issue item that served only to lend variety and to balance the USSR content of the test. Rather rigorous criteria were set up as the indication of emotional bias: The S had to accept the atmosphere conclusion on one USSR item and on the neutral item but reject the atmosphere conclusion on the opposite USSR item that disagreed with his or her attitude.

The number of responses for each type of conclusion (numbers 2 or 4; 1 or 3; 5) as well as the direction of response (proatmosphere, antiatmosphere, logical) was calculated. When chance expectancy and observed frequency were compared, the frequency of the 2 and 4 responses was much higher than chance expectancy. This was not unexpected because the 2 or 4 response is commonly chosen as the safest answer; it appears to be logically supported by the major and minor premises and at the same time has the loop-hole qualification "may" or "may not" which lures a S who is in doubt. Also, the 1 and 3 responses were chosen much less frequently than chance expectancy, and of the 102 responses, 86 were either pro- or anti-USSR, depending upon the direction of the atmosphere effect of the syllogism. Gorden (1953) viewed these 86 responses as simply an uncritical acceptance of the atmosphere effect rather than as conclusive indication of feeling on the issue.

Of the total 504 responses, 417 agreed with the atmosphere effect, 50 countered the direction of the atmosphere effect, and only 37 were the logically correct number 5. The nature of the 50 antiatmosphere conclusions were broken down into categories of neutral and USSR items. Forty of these antiatmosphere responses were made to USSR items while 10 were made to neutral items. The proportion of times the response countered the atmosphere effect was twice as great per item on the USSR items as on the neutral items. A further analysis indicated that of the 40 antiatmosphere responses to USSR items, 26 fulfilled the rigorous criteria for emotional bias, as opposed to a chance expectancy of 11 items. Gorden (1953) concluded that bias was present where the triplet criteria indicated it.

The finding of so few biases in this latter experiment may have provided insight into an artifact of the syllogistic reasoning task. The task is usually defined to Ss as a test of logical validity, and perhaps this forces Ss to consciously keep personal feeling out of their responses in most cases. Despite this possible artifact, however, Ss' beliefs and attitudes have been shown to be a definite factor that influences performance on syllogistic reasoning tasks.

GUIDELINES FOR SYLLOGISM CONSTRUCTION AND EXPERIMENTAL DESIGN

Despite the widespread use and study of syllogisms in psychology, concern about proper task construction appears to be lacking. Scarcely any literature provides guidelines on how to create a syllogism task in general; therefore, indirect means, such as rules and fallacies of syllogisms, serve as guidelines.

According to an introductory logic textbook (Copi, 1972), there are six rules for constructing standard-form syllogisms.

> Rule 1: A valid standard-form categorical syllogism must contain exactly three terms, each of which is used in the same sense throughout the argument.

The conclusion of a categorical syllogism asserts that a certain relation holds between two terms. The conclusion is justified only if the premises assert the relationship of each of the conclusion's terms to the same third term. Were these not asserted by the premises, no connection between the two terms of the conclusion would be established and the conclusion would not be implied by the premises. Therefore, three terms must be involved in every valid categorical syllogism. Any syllogism that contains more than three terms is invalid and is said to commit the fallacy of "four terms." If a term is used in different senses in the argument, it is being used equivocally; the fallacy of "equivocation" is committed.

The next two rules deal with distribution. A term is distributed in a premise when the premise refers to all members of the class designated by that term; otherwise the term is said to be undistributed in or by that proposition.

> Rule 2: In a valid standard-form categorical syllogism, the middle term must be distributed in at least one premise.

In the example "All dogs are mammals; all cats are mammals; therefore all cats are dogs," the middle term "mammals" is not distributed in either premise. Any syllogism that violates this rule is said to commit the fallacy of the "undistributed middle." At least one of the two conclusion terms must be related to the whole of the class designated by the third (middle) term; otherwise each may be connected with a different part of that class and thus not necessarily be connected with each other at all. In the preceding example, dogs and cats are included in part of the class of mammals. But different parts of that class are involved, so the middle term does not connect the syllogism's major and minor terms. For it to connect them, all the class designated by the middle term must be referred to in at least one premise.

Rule 3: In a valid standard-form categorical syllogism, no term can be distributed in the conclusion which is not distributed in the premises.

The conclusion of a valid argument cannot go beyond or assert any more than is contained in the premises. If the conclusion does go beyond what is asserted by the premises, the argument is invalid. A proposition that distributes one of its terms says more about the class designated by that term than it would if the term were undistributed. When the conclusion of a syllogism distributes a term that was undistributed in the premises, the conclusion says more about that term than the premises warrant; and the syllogism is therefore invalid. This invalidation can occur with either the major or the minor term, so Rule 3 may be broken in two ways.

When the major term of a syllogism is undistributed in the major premise but distributed in the conclusion, the fallacy of "illicit process of the major term" is committed. An example of this fallacy is "All dogs are mammals; no cats are dogs; therefore no cats are mammals." The conclusion makes an assertion about all mammals, saying that all of them are excluded from the class of cats. But the premises make no assertion about all mammals, so the conclusion illicitly goes beyond what the premises assert. When the minor term is undistributed in the minor premise but distributed in the conclusion, the argument commits the fallacy of "illicit process of the minor term." In the example "All communists are subversive elements; all communists are critics of the present administration; therefore all critics of the present administration are subversive elements," the premises make no assertion about all critics, but the conclusion does.

The next two rules are called rules of "quality" because they refer to the ways in which the negative quality of one or both premises restricts the kinds of conclusions that may validly be inferred.

Rule 4: No standard-form categorical syllogism that has two negative premises is valid.

Any negative proposition (E or O) denies class inclusion, asserting that all or some of one class is excluded from the whole of the other. Two negative premises can assert only that X is wholly or partially excluded from all or part of Y, and that Z is wholly or partially excluded from all or part of Y. But these conditions may very well obtain no matter how X and Z are related, whether by inclusion or exclusion, partial or complete. Therefore from two negative premises, no relationship whatever between X and Z can validly be inferred. Any syllogism which breaks Rule 4 commits the fallacy of "exclusive premises."

Rule 5: If either premise of a valid standard-form categorical syllogism is negative, the conclusion must be negative.

An affirmative conclusion asserts that one class is either wholly or partly contained in a second. This is justified only by premises which assert that a third class contains the first and is itself contained in the second.

In other words, both premises must assert class inclusion to entail an affirmative conclusion. But class inclusion can be expressed only by affirmative propositions. So an affirmative conclusion logically follows only from two affirmative premises. Thus, if either premise is negative, the conclusion cannot be affirmative but must be negative also. Any syllogism that breaks Rule 5 commits the fallacy of "drawing an affirmative conclusion from a negative premise."

The sixth and final rule concerns existential import. A proposition is said to have existential import if it asserts the existence of objects of some specified kind. For example, the premise "There are books on my desk", has existential import, whereas the proposition "There are no unicorns" does not. The particular propositions (I, O) have existential import, while the two universal propositions (A, E) do not.

Rule 6: No valid standard-form categorical syllogism with a particular conclusion can have two universal premises.

To break this rule is to go from premises that have no existential import to a conclusion that does. A particular proposition asserts the existence of objects of a specified kind, and to infer it from two universal premises, which do not assert the existence of anything at all, is clearly to go beyond what is warranted by the premises. For example, in the syllogism "All household pets are domestic animals; no unicorns are domestic animals; therefore some unicorns are not household pets," the conclusion asserts that there are unicorns whereas its premises do not. Any syllogism that violates Rule 6 may be said to commit the "existential" fallacy.

These six rules apply only to standard-form categorical syllogisms. (Similar rules regarding linear syllogisms are not apparent in the literature.) The rules provide an adequate test for the validity of any argument. If a standard-form categorical syllogism violates any of these rules, it is invalid; whereas if it conforms to all of them, it is valid. These rules should serve as guidelines when constructing a syllogistic reasoning task.

Johnson-Laird and Steedman (1978) suggest some basic requirements of an experiment that uses syllogisms as a task. The first suggestion is that Ss should have to make a deduction; that is, to create or produce a logically valid conclusion from the premises. If the task is merely to evaluate a given syllogism as valid or invalid, the S can carry it out without ever having to make an inference. Similarly, a multiple-choice task between commonly accepted conclusions may tend to obscure the deductive process, either because of the particular set of alternatives chosen by the experimenter or because of some idiosyncratic procedure that Ss adopt, such as backward processing or guessing. A second suggestion is that Ss should be given a representative selection of problems. Trying to draw general conclusions on the basis of perhaps 12 syllogisms (when 256 syllogisms are possible) gives a sense of futility. Another suggestion is that syllogisms should be presented with a sensible and noncontroversial linguistic content. Although there have been proponents for studying syllogistic reasoning with an abstract or symbolic

content, Wason and Johnson-Laird (1972) discovered that such syllogisms can lead to qualitative changes in performance. Syllogisms with abstract terms do not yield a purely deductive reasoning, uninfluenced by the subject's knowledge or attitudes, as much as they yield a kind of reasoning designed to compensate for the absence of everyday content. Finally, it is crucial to consider each syllogism separately in the analysis and description of results. Many researchers present only data pooled across different figures or across different moods. This may be appropriate for them in evaluating their own hypotheses, but it can render the data useless for anyone who wants to examine an alternative theory or to construct a general model of syllogistic reasoning.

REFERENCES

Begg, I.; and Denny, J. P. Empirical reconciliation of atmosphere and conversion interpretations of syllogistic reasoning errors. J Exp Psychol, 1969, 81(2):351-54.

Chapman, A. C.; and Chapman, J. P. Atmosphere effect reexamined. Psychol Rev, 1959, 58:220-26.

Clark, H. H. Linguistic processes in deductive reasoning. Psychol Rev, 1969, 76:387-404.

Copi, I. M. Introduction to logic. New York: MacMillan, 1972.

De Soto, C. B.; London, M.; and Handel, S. Social reasoning and spatial paralogic. Pers Soc Psychol, 1965, 2:513-21.

Dickstein, L. S. Conversion and possibility in syllogistic reasoning. Bull Psychonom Soc, 1981, 18(5):229-32.

Dickstein, L. S. Differential difficulty of categorical syllogisms. Bull Psychonom Soc, 1976, 8:330-32.

Dickstein, L. S. The effect of figure on syllogistic reasoning. Mem Cogn, 1978, 6(1):76-83.

Dickstein, L. S. Effects of instructions and premise order on errors in syllogistic reasoning. J Exp Psychol: Hum Learn Mem 1975, 1:376-84.

Erickson, J. R. Memory search in formal syllogistic reasoning. Paper presented at meeting of Midwestern Psychological Association, 1972.

Erickson, J. R. A set analysis theory of behavior in formal syllogistic reasoning tasks. In R. L. Solso (Ed.). Theories in cognitive psychology: The Loyola symposium. Hillsdale, NJ: Lawrence Erlbaum, 1974.

Frase, L. T. Associative factors in syllogistic reasoning. J Exp Psychol, 1968, 76(3):407-12.

Frase, L. T. Validity judgments of syllogisms in relation to two sets of terms. J Educ Psychol, 1966, 57:239-44.

Gorden, R. The effect of attitude toward Russia on logical reasoning. J Soc Psychol, 1953, 37:103-11.

Gottesman, L.; and Chapman, L. J. Syllogistic reasoning errors in schizophrenia. J Consult Psychol, 1960, 24:250-55.

Guyote, M. J.; and Sternberg, R. J. A transitive-chain theory of syllogistic reasoning (NR 150-412 ONR Technical Report No. 5). New Haven, CT: Department of Psychology, Yale University, 1978.

Henle, M. On the relation between logic and thinking. Psychol Rev, 1962, 69: 366-78.

Huttenlocher, J. Constructing spatial images: A strategy in reasoning. Psychol Rev, 1968, 75:550-60.

Janis, L.; and Frick, F. The relationship between attitudes toward conclusions and errors in judging logical validity of syllogisms. J Exp Psychol, 1943, 33:73-77.

Jenkins, J. J. Mediated associations: Paradigms and situations. In C. Cofer and M. Musgrave (Eds.). Verbal behavior and learning. New York: McGraw, 1963.

Johnson-Laird, P. N.; and Steedman, M. The psychology of syllogisms. Cogn Psychol, 1978, 10:64-99.

Kaufman, H.; and Goldstein, S. The effects of emotional value of conclusions upon distortions in syllogistic reasoning. Psychonom Sci, 1967, 7:367-68.

McGuire, W. J. Cognitive consistency and attitude change. Abnorm Soc Psychol, 1960, 60:345-53.

Morgan, J. J. B.; and Morton, J. T. The distortion of syllogistic reasoning produced by personal convictions. J Soc Psychol, 1944, 20:39-59.

Quinton, G.; and Fellows, B. "Perceptual" strategies in the solving of three-term series problems. Br J Psychol, 1975, 66:69-78.

Revlin, R.; Ammerman, K.; Petersen, K.; and Leirer, V. Category relations in syllogistic reasoning. J Exp Psychol, 1978, 70(4):613-25.

Revlis, R. Two models of syllogistic reasoning: Feature selection and conversion. J Verb Learn Verb Behav, 1975, 14:180-95.

Roberge, J. J. Further examination of mediated asssociations in deductive reasoning. J Exp Psychol, 1971, 87(1):127-29.

Sells, S. B. The atmosphere effect: An experimental study of reasoning. Arch Psychol, 1936, No. 200.

Simpson, M. E.; and Johnson, D. M. Atmosphere and conversion errors in syllogistic reasoning. J Exp Psychol, 1966, 72(2):197-200.

Sternberg, R. J. Reasoning with determinate and indeterminate linear syllogisms. Br J Psychol, 1981, 72:407-20.

Sternberg, R. J.; and Turner, M. E. Components of syllogistic reasoning. Acta Psychol, 1981, 47:245-65.

Sternberg, R. J.; and Weil, E. M. An aptitude x strategy interaction in linear syllogistic reasoning. J Educ Psychol, 1980, 72(2):226-39.

Thurstone, L. L. Primary mental abilities. Psychometric Monograph, no. 1. Chicago: University of Chicago Press, 1938.

Von Domarus, E. The specific laws of logic in schizophrenia. In J. S. Kasanin (Ed.). Language and thought in schizophrenia. Berkeley: University of California Press, 1944.

Wason, P. C.; and Johnson-Laird, P. N. Psychology of reasoning: Structure and content. London: Batsford, 1972.

Woodworth, R. S.; and Sells, S. B. An atmosphere effect in syllogistic reasoning. J Exp Psychol, 1935, 18:451-60.

END

FILMED

11-85

DTIC